14:38:59 OCA PAD AMENDMENT - PROJECT HEADER INFORMATION 10/06/94

Active

Project #: E-24-603 Cost share #: Rev #: 6
Center # : 10/24-6-R7664-0A0 Center shr #: OCA file #:
Work type : RES
Contract#: 5-24209 Mod #: LTR DTD 9/19/94 Document : AGR
Prime #: NCR-9206764 Contract entity: GTRC

Subprojects ? : Y CFDA: 47.070
Main project #: PE #: N/A

Project unit: ISYE Unit code: 02.010.124
Project director(s):
CLARKE L W ISYE (404)894-2325
NEMHAUSER G L ISYE (404)-

Sponsor/division names: UNIVERSITY OF PENNSYLVANIA / PHILADELPHIA, PA
Sponsor/division codes: 400 / 137

Award period: 920901 to 960229 (performance) 960531 (reports)

Sponsor amount	New this change	Total to date
Contract value	10,002.00	130,266.00
Funded	52,320.00	130,266.00
Cost sharing amount		0.00

Does subcontracting plan apply ?: N

Title: DESIGN OF SURVIVABLE HETEROGENEOUS DATA NETWORKS

PROJECT ADMINISTRATION DATA

OCA contact: Ina R. Lashley 894-4820

Sponsor technical contact

DR. G. ANANDALINGAM
(000)000-0000

DEPARTMENT OF SYSTEMS
UNIVERSITY OF PENNSYLVANIA
PHILADELPHIA, PA 19104

Sponsor issuing office

ANTHONY MERRITT, DIRECTOR
(215)898-7293

OFFICE OF RESEARCH ADMINISTRATION
UNIVERSITY OF PENNSYLVANIA
133 SOUTH 36TH STREET, SUITE 300
PHILADELPHIA, PA 19104-3246

Security class (U,C,S,TS) : U
Defense priority rating : N/A
Equipment title vests with: Sponsor
NONE PROPOSED.

ONR resident rep. is ACO (Y/N): N
NSF supplemental sheet
GIT X

Administrative comments -
LTR DTD 9/19/94 ADDS FINAL INCREMENT OF $52,320 AND EXTENDS PERIOD OF PERFORMANCE ONE YEAR. ($10K "PARTICIPANT SUPPORT" TRANSFERRED TO SUB-PROJECT)

GEORGIA INSTITUTE OF TECHNOLOGY
OFFICE OF CONTRACT ADMINISTRATION

U

NOTICE OF PROJECT CLOSEOUT

Closeout Notice Date 02/29/96

Project No. E-24-603 Center No. 10/24-6-R7664-0A0

Project Director CLARKE L W School/Lab ISYE

Sponsor UNIVERSITY OF PENNSYLVANIA/PHILADELPHIA, PA

Contract/Grant No. 5-24209 Contract Entity GTRC

Prime Contract No. NCR-9206764

Title DESIGN OF SURVIVABLE HETEROGENEOUS DATA NETWORKS

Effective Completion Date 960229 (Performance) 960531 (Reports)

Closeout Actions Required:	Y/N	Date Submitted
Final Invoice or Copy of Final Invoice	Y	
Final Report of Inventions and/or Subcontracts	Y	
Government Property Inventory & Related Certificate	N	
Classified Material Certificate	N	
Release and Assignment	N	
Other	N	

Comments

Subproject Under Main Project No.

Continues Project No.

Distribution Required:

Project Director	Y
Administrative Network Representative	Y
GTRI Accounting/Grants and Contracts	Y
Procurement/Supply Services	Y
Research Property Managment	Y
Research Security Services	N
Reports Coordinator (OCA)	Y
GTRC	Y
Project File	Y
Other	N
	N

NOTE: Final Patent Questionnaire sent to PDPI.

E24-603
#1

Annual Progress Report for Design of Survivable Heterogeneous Data Networks

NSF Grant #

Lloyd W. Clarke

Overview of Research

We began work on the above NSF grant in September, 1992. this work involves the development of novel solution techniques for designing networks to survive under link or node failure. The interconnection of Local Area Networks, the connection of LANs to Wide Area Networks, the interconnection of WANs, are all examples of heterogeneous data networks. Each of these networking problems have unique characteristics based on the communications protocols and the routing behavior of the interconnect devices.

The first half of the research is focused on mathematical aspects of appropriate network models. Since one of the major issues of LAN and WAN networks is data flow, we concentrate on a network model that uses a path oriented design approach. The path approach can guarantee transmission from origin to destination of a specified amount of data.

The second half of the research will apply and expand these algorithms to specific scenarios from the telecommunications industry. We are currently pursuing a working relationship with companies in the telecommunications industry that would be capable and interested in a information exchange.

Below we summarize our progress in the area of network algorithms.

Progress for Network Algorithms

Our work began with the investigation of the link based network algorithms. We make the distinction between network design based on links where one has to decide what *links* to place in the network to ensure survivability, and design based on *paths* where one has to choose among a set of paths for each origin-destination pair to ensure survivability. The link based approach can be done efficiently in many instances. Work done in this area is described in the enclosed paper "A System of Designing Minimum Cost Survivable Networks," written by the two P.I. and submitted for publication. The paper includes the following:

- Formulations for link survivable and node survivable models.
- Methods for finding the lower bound and optimal solution via mathematical programming
- Design heuristics for an initial survivable network and ways to improve the design.
- Several techniques are put together for a comprehensive system of design, with numerical results.

In the progression from link based formulation to path based formulation we developed a paper that provides the reader with tools and background information needed to do research in the area of survivable network design. The paper "The Design of Minimum Cost Survivable Telecommunication Networks: A Synthesis" (enclosed) is still in draft form. The paper covers:

- Brief description of network reliability
- Fundamental concepts of network survivability, disjoint paths, and biconnected subgraphs
- Formulations and sample research using link based designs
- Formulations and sample research using path based designs

The current focus of our research is finding an efficient solution for the path-based survivable network formulation. The path based approach in some respects is difficult to deal with than the link based approach. The path based approach has an exponential number of variables. This can be dealt with using a column generation approach. Research in the area of column generation is not as mature as cutting planes (used for link based approaches). We have a formulation for the path based approach. Panjing Gong, Ph.D. student of Lloyd Clarke, has developed a preliminary version of the code to solve this problem using column generation. We have been invited to discuss our work in the Fall '93 ORSA/TIMS Joint National Meeting in Phoenix.

E-24-603
2

Annual Progress Report for Design of Survivable Heterogeneous Data Networks

NSF Grant NCR 92-06764
(September 1993 to August 1994)

G. Anandalingam, University of Pennsylvania
Lloyd W. Clarke, Georgia Institute of Technology

This is the second year of research on the above NSF grant. The work involves the development of formulations and novel solution techniques for designing heterogeneous data networks to survive under link or node failure.

1 Survivable Network Algorithms

We made substantial progress on this research. We have completed much of the work on link based network algorithms, and have produced a software package that, when exercised, yields a survivable topology. This software has been used in our research on international networks and in LAN-LAN networks (both to be described later). As outlined in our Progress Report last year (July 1993), this work involves formulating the survivable network design problem, obtaining heuristics for solving problems with numerous nodes, integrating these heuristics into a Design System, and obtaining numerical results that show the computational performance of the Design System. We have done all of these things.

The highlights on this research was the final revisions made to a paper that we (Lloyd Clarke and G. Anandalingam) submitted to the journal *Computers and Operations Research* entitled "A Bootstrap Heuristic for Designing Minimum Cost Survivable Networks"; this paper should appear in print shortly. In February 1994 we have submitted another paper entitled "An Integrated System for Designing Minimum Cost Survivable Telecommunications Networks" for publication. We expect to hear a decision on this shortly. We also have a Survey Paper entitled "The Design of Minimum Cost Survivable Telecommunications Networks: A Synthesis" that provides the reader with tools and background information needed to do research in this area. This paper has been circulated for comments, and is being continuously revised. We hope to complete this process soon, and then submit the paper for publication.

Early portion of this work emphasized link based network algorithms. The current emphasis is path based algorithms. We make the distinction between network design based on links where one has to decide what *links* to place in the network to ensure survivability, and design based on *paths* where one has to choose among a set of paths for each origin-destination pair to ensure survivability. The path based approach has the advantage of better being able to deal with capacity constraints and origin destination flow requirements.

The path based algorithm is the basis of the dissertation of Panjing Gong, a Ph.D. student supervised by Lloyd Clarke at Georgia Institute of Technology. With the path based models, the number of variables is a factorial function. A 10 node network can have over 2 million variables. This is handled by using column generation techniques, where the initial problem will only have a small subset of the variables and additional variables are added as needed as the algorithm progresses.

Determining how and when to add variables to the model requires a reasonable amount of optimization theory. This theory isn't easily applied to survivable networks. The current emphasis is general networks with capacity constraints. The available capacity is modeled as DS0 and DS1 lines. There is a choice of a discrete number of each of these two line types. Network design problems with discrete capacity choices are very difficult problems. Optimal solutions can be found for problems in size of up to 20 nodes. Comparing our path based formulation to a standard arc based formulation we can find optimal solutions 5 to 10 times faster. These results were presented at the Fall '93 ORSA/TIMS Joint National Meeting in Phoenix and the Institute of Industrial Engineers (IIE) Research Conference in Atlanta, March 1994. A copy of the presentation slides are enclosed.

The path based formulation now includes three models to represent different methods of combining the different capacity lines available. We are beginning to address the routing issue. Different types of networks will allow flow for a specific origin destination pair to travel multiple routes. The one route case, adds additional integer variables and significantly complicates the column generation. The additional models and routing constraints is the current emphasis of the path algorithm research. We are seeking feed back from the optimization community by presenting our work in the 15th International Symposium on Mathematical Programming in Ann Arbor, August 1994. We have accepted an invitation from the ORSA Technical Section on Telecommunications to present this work at the Fall '94 ORSA/TIMS Joint National Meeting in Detroit and hope to get useful feedback from this group. In addition to conference presentations, we are currently preparing a paper and Ph.D. proposal.

2 Design of Survivable International Networks

This research is the basis of the dissertation of Keesung Nam, a Ph.D. student supervised by G. Anandalingam at the University of Pennsylvania entitled *The Design of Survivable International Telecommunications Networks* (Table of Contents attached). The main differences between designing survivable data networks for corporations located in one country, and distributed across the globe are: (a) the peak communications traffic between pairs of countries occurs at different times, and (b) different countries have different objectives, although they may use the same network. Thus, telecommunications networks can be used very efficiently by implementing dynamic routing that sends traffic via transit countries.

In Nam's dissertation, we have formulated two kinds of problems. First, we have examined a survivable network that uses pre-planned time varying dynamic routing based on the daily traffic profile among countries (Chapter 3). Each country-to-country pair can choose the degree of survivability by giving the network designer a parameter that fixes the amount of traffic that should be re-routed under node or link failure. We have also derived heuristics to solve large size problems and have tested the heuristics against lower bounds. The heuristic is very simple: It involves decomposing the problem, using the Design System of Clarke and Anandalingam to solve part of the new structure, and in the other part, using linear programming to solve integer programming problems and rounding the solutions in a pre-arranged order. We have tested the problem on both an Atlantic and a Pacific Basin network. The Pacific Basin network is large (containing 15 countries) with more than 40,000 variables, many of them integer, and around 60,000 constraints. The problem was solved in 16-33 minutes on a SUN SparcStation, and was within 1 per cent of the Lower Bound. These are excellent results.

The second problem we examined involved the advantages and disadvantages for the countries to form a coalition in order to design a truly integrated network. The issues were the increase in benefits and decrease in costs to the entire coalition of designing the network cooperatively, and how these benefits should be distributed to the players equitably. (Chapter 5 in Nam's dissertation). For this we used developments in Cooperative Game Theory. We also wanted to examine whether the U. S. should take the lead in developing the international network. The issue here is whether there are economic gains to be made by being the leader in a non-cooperative Stackelberg game, or whether one should only be part of a coalition. To analyze this case, we had to develop some methodology in bi-level programming (Chapter 4). All of the numerical analysis in these cases were based on real world data that was distilled very carefully from published sources.

We will be presenting a paper on this research at a Workshop in the United Kingdom, and hope to submit two papers on this research to telecommunications journals.

3 The LAN-LAN Internetworking Problem

We have made significant progress in this area of research which will form much of the Ph.D. dissertation of Frances Lee who is supervised by G. Anandalingam at the University of Pennsylvania. This work will continue on to next year, and we expect that the dissertation will be completed by May 1995.

We have examined a number of bridges and routers in the commercial market, and have extracted out all the important features. The issue is which device one would use in interconnecting different LANs. This examination has resulted in a draft of a paper entitled "On the Choice of Bridges and Routers as Internetworking Devices". *We want to emphasize that this paper is a preliminary draft and is expected to undergo substantial revisions during the summer, July - August 1994.*

We have also modeled the survivable LAN-LAN internetworking problem using a hierarchical architecture. In this case, we use either a combination of remote routers and backbone routers or bridges and backbone routers. The remote network is designed using both a mathematical programming approach and a statistical clustering approach. We have examined the computational performance of each. The backbone network is designed using the Design System that was produced by Clarke and Anandalingam, described earlier. We also have a heuristic for connecting the remote and the backbone networks. The methodology has been applied to the real world case of designing a LAN-LAN network for the Penn campus. The results are quite encouraging. We have produced a paper entitled "Design of Survivable LAN-LAN Internetworks with a Hierarchical Topology". *Again, we have to caution the reader that this is a preliminary draft that is expected to undergo substantial revisions during the Summer, July - August 1994.*

We will enhance the work described in the preceding paragraph by applying the heuristics developed to problems with many nodes and providing extensive numerical tests. In addition, we need to provide dynamic routing schemes for end-to-end traffic in a LAN-LAN internetwork. We will also extend our analyses to LAN-WAN and WAN-WAN internetworks where we will use a combination of routers and gateways. We expect to make modifications to the previous work done by Clarke and Anandalingam on developing a Design System for survivable networks. The functionality of routers and gateways make it imperative that we provide some fresh ideas to the earlier work.

4 Availability of Self-Healing Ring Networks

We have also been working with Mark Wilson of AT&T Bell Laboratories on estimating the availability of self healing ring networks which will be his Ph.D. dissertation at the University of Pennsylvania. His dissertation proposal has also been enclosed. While Mark Wilson is not being funded by the NSF grant, his dissertation is intimately connected to the subject of survivable heterogeneous networks.

Mark Wilson is examining models for evaluating the availability of networks that are supposed to be survivable (or self healing). While much of our previous work has been on providing a survivable topology, Wilson's dissertation is on evaluating the probabilities of such networks actually being survivable at points in time and after some failure event (i.e. the 'availability' of such networks). The focus of this work will be on SONET (Synchronous Optical Network) which is a survivable network because of its ring topology, and 'availability' will be one of components that customers will use (along with cost, functionality etc) in order to assess their networks.

The collaboration will involve the development of heuristics to estimate availability rapidly. In addition, once these availability models are calibrated on SONET, we expect that they will be used on the topologies that we produce in our work described above. Wilson's research is ongoing and will be partially supervised by G. Anandalingam. Mark Wilson also provides some very good insights from an industry perspective on our ongoing research efforts.

E-24-603 #3

Annual Report for Design of Survivable Heterogeneous Data Networks

NSF Grant NCR 92-06764
(June 1994 through May 1995)

Lloyd W. Clarke
Industrial & Systems Engineering
Georgia Institute of Technology

Overview of Research

This is the final year of research on the above NSF grant. The work involves the development of formulations and novel solution techniques for designing heterogeneous data networks to survive under link or node failure. The interconnection of Local Area Networks, the connection of LANs to Wide Area Networks, the interconnection of WANs, are all examples of heterogeneous data networks. Each of these networking problems have unique characteristics based on the communications protocols and the routing behavior of the interconnect devices.

We have concentrated on network algorithms to solve these problems. We will describe two areas on which we have concentrated. The first area is heuristic algorithms for survivable networks. The second area is mathematical algorithms for general network design.

Survivable Algorithms

Our research in the area of survivable networks has included both heuristics and mathematical programming approaches. We have developed a software package that designs both capacitated and uncapacitated networks. The heuristic portion of the system works in two steps. The first step is the design of an initial feasible survivable network. The

1 Introduction

This paper addresses the discrete choice network design problem, which arises in the telecommunication industry. Among the problems in telecommunication industry which are related to network design, one is the private network leasing problem. Instead of using the public switched network, large organizations can lease private lines to use exclusively to carry its communication traffic between different locations. Because these private lines are charged on a fixed rate, the overall cost might be less for an organization to lease a private line between any two locations with heavy communication traffic than to use the public switched network and pay the bill on a per usage basis. Another advantage for an organization to lease the private network is that the organization has more flexibility to reconfigure the network to accomodate communication traffic changes in the organization. In addition, private networks offer more reliable service.

An organization can lease exclusive lines from its offices to the nearest telephone company central office and exclusive lines between the central offices that connect different locations of the organization. There are some facilities with different capacities to be chosed from for the private networks. For example, DS0(Digital Signal Level 0), DS1(Digital Signal Level 1), which has the capacity equivalent to 24 DS0 facilities, and DS3(Digital Signal Level 3), which has the capacity equivalent to 28 DS2 facilities. A DS1 circuit may cost the same as 8 to 10 DS0 circuit, which depends on the length of the circuit.

Given an organization's forecast for the communication traffic between its offices in various locations, what configuration of the facilities between the central offices should be leased to carry this traffic at minimum cost? The cost for the private network corresponds to the leasing cost of the facilities loaded on the links. There is no routing cost for the user.

1.1 Formulation

We can formulate the private network leasing problem as mixed integer programming problem. The basic ingredients of the model are a set V of nodes and a set E of links that are available for designing the network (we will specify whether the network is directed or not for each model). We also use the following notations: T is the number of different types of facilities; c_r is the capacity of facility r ($1 \leq r \leq T$); k_{ij}^r is the cost of facility r on link (i,j). This model permits multiple commodities. Let $H = \{h = (s,t) \mid (s,t) \text{ is a source} - \text{terminal pair}\}$. For each $h \in H$, let R_h denote the requirement between the source, denoted s(h), and the terminal, denoted t(h).

Two types of variables are included in this model, one for discrete choice design decisions, one for continuous flow decisions. Let y_{ij}^r be integral variable that indicates how many facility r are loaded on link (i,j). Let f_{ij}^h denote the flow of commodity h on link (i,j). The general model on a directed network is:

In Model 1, constraints (1) are the flow conservation constraints, which correspond to the flow conservation for each commodities at each node; constraints (2) are the capacity constraints, which force the requirement that the total flow on any link can not exceed the

We make the distinction between network design based on links where one has to decide what *links* to place in the network to ensure survivability, and design based on *paths* where one has to choose among a set of paths for each origin-destination pair to ensure survivability. The path based approach has the advantage of better being able to deal with capacity constraints, origin destination flow requirements, and delay constraints.

The path based models require a completely different mathematical approach. To better develop this approach, we return to the basic network design problem without survivability constraints. In these models, the number of variables is a factorial function. A 10 node network can have over 2 million variables. This is handled by using column generation techniques, where the initial problem will only have a small subset of the variables and additional variables are added as needed as the algorithm progresses. Determining how and when to add variables to the model requires a reasonable amount of optimization theory. This theory isn't easily applied to survivable networks. The emphasis is general networks with capacity constraints.

We worked with three different models. The first model is a simple capacitated model with choice of link locations. The second model has multiple size links available. Each location can have only one of the several different size links. The final model has multiple link sizes and can install any number of links in any location. Each of these models have a split flow and a unified flow version. To work with the requirement that origin-destination flow must remain on one path, required some innovative mathematical thought that relates to the column generation in branch and bound. We have found that the path formulation method is a more efficient and very promising approach to solving network design problems. Optimal solutions can be found for problems in size of up to 20 nodes. Comparing our path based formulation to a standard arc based formulation we can find optimal solutions up to 5 times faster.

We have presented this research in its different stages at different technical conferences. These presentations include the

- Fall '94 ORSA/TIMS Joint National Meeting in Detroit
- 15^{th} International Symposium on Mathematical Programming in Ann Arbor, August 1994
- Institute of Industrial Engineers (IIE) Research Conference in Atlanta, March 1994
- Fall '93 ORSA/TIMS Joint National Meeting in Phoenix

network design problems. For example, if we add survivability constraints to our design problem, it is much easier to deal with the extra survivability constraints with the path-based model than with the link-based model. Although we do not study the survivability problem in this paper, it was the inspiration for the present work. In section 2.3, we will see that there are less rows but much more columns in the path-based model than in the link-based model. Therefore, a branch-and-price algorithm is introduced to solve the path-based model. In order to show the effectiveness of our branch-and-price algorithm, we compare it to a branch-and-bound algorithm to solve the link-based model. Computational results on the two different algorithms are compared.

In Model 1, if we change the flow variables f_{ij}^h to be binary, we have Model 2, which is a much harder problem than Model 1. Model 2 is very practical, since for some real problems, the commodities are unsplitable such as phone calls.

We will use similar strategies to solve Model 2 as we do on Model 1. Both link-based model and path-based model are studied, and branch-and-price algorithm is used to solve the path-based model. For the branch-and-price algorithm on path-based model, since the algorithm will branch on some binary path (flow) variables, some new branching rule must be given to deal with it. In Section 5, we discuss our algorithms on Model 2 based on Model 1. Two branching rules for the branch-and-price algorithm are given on solving the path-based model. The computational results on the two algorithms with different branching rules, as well as the computational results on the link-based model, are campared.

2 Formulations

In this paper, we will discuss some special versions of the network design problem. Based on Model 1, the general model, we make the following assumptions:

(1) No flow costs, i.e. $b_{ij}^h = 0$.

Under these assumptions, three models which are slightly different only on the constraint how the facilities can be loaded on a link are studied. In section 2.1, we set up the link-based models. In section 2.2, we set up the path-based models.

2.1 Link-based models

For the link-based models, we assume that the network is directed. All variables are defined as described in the previous section.

Model (L_1):

Under the assumptions that only high capacity facility is used (i.e. T=1, c=high capacity) and at most one facility is loaded for any link (i.e. y_{ij} is binary), we have the model(L_1).

In Model L_1, (5) is the flow conservation constraint, (6) is the capacity constraint.

Model (L_2):

Under the assumptions that T different capacity facilities are used and at most one facility is loaded for any link (i.e. y_{ij}^r is binary), we have the model(L_2).

Model (L_3):

$$\min \sum_{(i,j)\in E} k_{ij} y_{ij}$$

$$\begin{aligned}
\sum_{j\in V,(i,j)\in E} f_{ij}^h - \sum_{l\in V,(l,i)\in E} f_{li}^h &= \begin{cases} R_h, & \text{if } i = s(h) \\ -R_h, & \text{if } i = t(h) \\ 0, & \text{o/w} \end{cases} && \forall i \in V, \forall h \in H. && (5)\\
\sum_{h\in H} (f_{ij}^h + f_{ji}^h) &\leq c y_{ij} && \forall (i,j) \in E && (6)\\
y_{ij} \ & binary && \forall (i,j) \in E && (7)\\
f_{ij}^h &\geq 0 && \forall (i,j) \in E, h \in H && (8)
\end{aligned}$$

Model L_1: Link-based model with only high capacity facility and binary y's

$$\min \sum_{(i,j)\in E} \sum_{r=1}^{T} k_{ij}^r y_{ij}^r$$

$$\begin{aligned}
\sum_{j\in V,(i,j)\in E} f_{ij}^h - \sum_{l\in V,(l,i)\in E} f_{li}^h &= \begin{cases} R_h, & \text{if } i = s(h) \\ -R_h, & \text{if } i = t(h) \\ 0, & \text{o/w} \end{cases} && \forall i \in V, \forall h \in H. && (9)\\
\sum_{h\in H} (f_{ij}^h + f_{ji}^h) &\leq \textstyle\sum_{r=1}^{T} c_r y_{ij}^r && \forall (i,j) \in E && (10)\\
\sum_{r=1}^{T} y_m^r &\leq 1 && \forall e_m \in E && (11)\\
y_{ij}^r \ & binary && \forall (i,j) \in E, 1 \leq r \leq T && (12)\\
f_{ij}^h &\geq 0 && \forall (i,j) \in E, h \in H && (13)
\end{aligned}$$

Model L_2: Link-based model with Binary y's

If we allow more than one facilities to be loaded on each link, then change y variables to integers and delete constraints (11) in model (L_2), we get model (L_3).

2.2 Path-based models

We now set up the alternative models for the network design problems we study, and refer them as path-based models. For the path-based models, we assume that the network is undirected.

Let

$P^h = \{p_1^h, \ldots, p_{n_h}^h\}$ be the set of all paths joining $s(h)$ and $t(h)$ for $h \in H$

and

$$a_{mj}^h = \begin{cases} 1, & if\ e_m \in p_j^h\ for\ m = 1, \ldots, M,\ 1 \leq h \leq |H|, \\ & 1 \leq j \leq n_h \\ 0 & o/w \end{cases}$$

and

x_j^h = flow of commodity h through path p_j^h

Then, the path-based model corresponding to the link-based model(L_1) is:

Model (P_1):

$$\min \sum_{e_m \in E} k_m y_m$$

$$\sum_{h \in H} \sum_{j=1}^{n_h} a_{mj}^h x_j^h \leq c y_m \quad \forall e_m \in E \qquad (14)$$

$$\sum_{j=1}^{n_h} x_j^h = R_h \quad \forall h \in H \qquad (15)$$

$$y_m\ binary \quad \forall e_m \in E \qquad (16)$$

$$x_j^h \geq 0 \quad \forall h \in H, 1 \leq j \leq n_h \qquad (17)$$

Model P_1: Path-based model with Binary y's

In Model P_1, (14) is the capacity constraint, (15) is the flow conservation constraint.

If we allow more than one facilities to be loaded on each link, then change y variables to integers and delete constraints (20) in model (P_2), we get model (P_3).

Model (P_2):

$$\min \sum_{e_m \in E} \sum_{r=1}^{T} k_m^r y_m^r$$

$$\begin{aligned}
\sum_{h \in H} \sum_{j=1}^{n_h} a_{mj}^h x_j^h &\leq \textstyle\sum_{r=1}^{T} c_r y_m^r && \forall e_m \in E && (18)\\
\sum_{j=1}^{n_h} x_j^h &= R_h && \forall h \in H && (19)\\
\sum_{r=1}^{T} y_m^r &\leq 1 && \forall e_m \in E && (20)\\
y_m^r \ &binary && \forall e_m \in E, 1 \leq r \leq T && (21)\\
x_j^h &\geq 0 && \forall h \in H, 1 \leq j \leq n_h && (22)
\end{aligned}$$

Model P_2: Path-based Model with Binary y's

2.3 Size Comparison

The sizes of the link-based model(L_2) and the path-based model(P_2) are compared in the following tables.

	# of rows	# of columns
model (L_2)	$\|V\|\|H\| + 2\|E\|$	$\|E\|(\|H\| + T)$
model (P_2)	$\|H\| + 2\|E\|$	$T\|E\| + g(\|E\|)$

Table 1: Comparison of model (L_2) and (P_2)

Where g() is an exponential function. Let $n = |V|$, for a complete graph, $g(n) = C_n^2 \sum_{k=0}^{n-2} C_{n-2}^k k! > n!$

	# of rows	# of columns
model (L_2)	540	2115
model (P_2)	135	$> 2,000,000$

Table 2: On a 10-node complete graph

Note: Suppose $|H| = |E|$ and T=2.

3 Solution Methods

In Table 2, we see that even in a 10-node complete graph, the number of columns in model(P_2) is very big. Therefore, it is very difficult, if not impossible, to solve the relaxed LP of the path-based model directly. In this section, we give the branch-and-price algorithm to solve the path-based model(P_2). In this algorithm, instead of solving the relaxed LP of path-based model(P_2) directly, column generation technique is used to do it. The proof of the correctness of the algorithm is also presented. This algorithm will also work for path-based model(P_1) and model(P_3) too.

3.1 Column Generation Algorithm

Branch-and-price algorithm will be used to solve the path-based model(P_2). It means that at every node of the branch-tree, column generation algorithm will be used to solve the relaxed LP.

In general, the column generation algorithm is used to solve large-scale linear programs or linear programs that contain constraints of special structure. The strategy of the column generation algorithm is to operate on two separate linear programs, one is the master problem, the other one is the subproblem. After the master problem is solved, a new set of cost coefficients to the subproblem is generated and passed to the subproblem; then the subproblem with the new cost coefficients is solved, and one or more new columns for the master problem are generated, which are added to the master problem to form a new master problem. Repeat the above procedure, until a point is reached where the solution to the original problem is achieved.

There are two kinds of suitable LP models for column generation algorithm. One is the decomposable model, in which the constraints of the LP are divided into two sets: general constraints and constraints with special structure. For the decomposable model, master problem is an LP over the set of general constraints, and the subproblem is an LP over the set of special constraints. The other one is the nondecomposable model, which includes nondecomposable LP problems with too many variables. For nondecomposable model, the master problem is an LP with a subset of all the variables, and the subproblem is to price out all the variables not included in the master problem.

The LP relaxation of our path-based models belongs to the nondecomposable model. In section 3.1.1, we will present the column generation algorithm for the LP relaxation of path-based model(P_2). For path-based model(P_1) and model(P_3), it works too. It turns out that the subproblem is induced to the shortest path problem.

3.1.1 The master problem and the subproblem

After relaxing the binary y variables in model(P_2) to be continuous and introducing artificial variables s_h for $h \in H$, we have the relaxed LP of model(P_2) with all the possible columns as:

(PLP): The LP relaxation of (P_2)

$$\min \sum_{e_m \in E} \sum_{r=1}^{T} k_m^r y_m^r + \sum_{h \in H} A s_h$$

$$\begin{array}{llcll}
w_m: & \sum_{r=1}^{T} c_r y_m^r - \sum_{h \in H} \sum_{j=1}^{n_h} a_{mj}^h x_j^h & \geq 0 & \forall e_m \in E & (23) \\
v_h: & \sum_{j=1}^{n_h} x_j^h + s_h & = 1 & \forall h \in H & (24) \\
u_m: & \sum_{r=1}^{T} y_m^r & \leq 1 & \forall e_m \in E & (25) \\
 & y_m^r & \geq 0 & \forall e_m \in E, 1 \leq r \leq T & (26) \\
 & x_j^h & \geq 0 & \forall h \in H, 1 \leq j \leq n_h & (27) \\
 & s_h & \geq 0 & \forall h \in H & (28)
\end{array}$$

Where w_m, v_h and u_m are the corresponding dual variables, and A is a big number serving as the penalty coefficient for the artificial variables s_h.

In (PLP), we see that there are two kinds of variables: the y variables and the x variables. The number of y variables is much smaller than the number of x variables. We will include all the y variables in the initial master problem, and implement column generation algorithm only on x variables.

Suppose the x variables corresponding to paths in the set MP^h, $\forall h \in H$ are included in the master problem (For the initial master problem, we can set $MP^h = \emptyset$, $\forall h \in H$, and include the artificial variable s_h for each $h \in H$). Then, after we solve the master problem, we need to price out the x variables not included in the master problem. Let (x^*, y^*, s^*) denote the optimal solution of the master problem, from (PLP), the price-out problem is to check if

$$(Inequality\ 1) \qquad v_h^* \leq \sum_{m=1}^{M} a_{mj}^h w_m^* \qquad \forall h \in H,\ p_j \in P^h - MP^h \qquad (Condition\ 1)$$

where $w_m^* \geq 0$.

Let $w_m^* \geq 0$ be the weight for $e_m \in E$ in the network. If $(Condition\ 1)$ satisfied, then (x^*, y^*, s^*) is the optimal solution of (PLP); Otherwise, any path variable corresponding to a path which violates $(Inequality\ 1)$ can be added to the master problem.

Let p^h be the shortest path joining s(h) and t(h) in the network, then $(Condition\ 1)$ is equivalent to $(Condition\ 2)$:

$$v_h^* \leq \sum_{e_m \in p^h} w_m^* \qquad \forall\ h \in H \quad (Conditon\ 2)$$

If (*Condition* 2) satisfied, then (x^*, y^*, s^*) is the optimal solution of (PLP); Otherwise, the path variables corresponding to the shortest paths that violate (*Conditon* 2) will be added to the master problem.

Therefore, the subproblem of our column generation algorithm is induced to a shortest path problem, which is:

(SP): Subproblem

Solve the shortest path problem for all $h \in H$ in the network with weight $w_m^* \geq 0$ for $e_m \in E$.

Note: We can use Floyd's Algorithm to solve the shortest path problem for all $h \in H$.

3.1.2 The Column Generation Algorithm and the proof of its correctness

Algorithm to solve (PLP):

Step 0. Set (M_0) to be the initial master problem, $i = 0$;

Step 1. Solve (M_i), get dual variables w_m^* and v_h^*;

Step 2. Solve (SP);

Step 3. If (Condition 2) satisfied, stop, the optimal solution of (M_i) is the optimal solution of (PLP);
Otherwise, let

$$PATH = \{p^h \mid v_h^* > \sum_{e_m \in p^h} w_m^*\}$$

add columns corresponding to $p^h \in PATH$ to (M_i) to form (M_{i+1}), set $i = i + 1$, goto step 1.

Proof of the correctness of the algorithm:

In step 3, we see that every column added to the model is a column of (PLP). Since the number of all possible columns for (PLP) is limited, this algorithm will terminate in finite step with a solution (x^*, y^*, s^*). Since (condition 2) is satisfied when the algorithm terminates, (x^*, y^*, s^*) is the optimal solution of (PLP).

In the optimal solution (x^*, y^*, s^*) of (PLP), if s^* is not a zero vector then (PLP) is infeasible, so is the integer model(P_2), stop.

3.2 Branch-and-price algorithm to solve model(P_2)

We will use branch-and-price algorithm to solve model(P_2).

In section 3.1, we studied the column generation algorithm for solving the relaxed LP (PLP) at the root node of the branch-tree. One thing we have observed is that at the root node we include all the integer y variables in our initial model, and we only generate new columns corresponding to the continuous path variable x. Therefore, at any node of the branch-tree with some y variables being fixed, we can use the column generation algorithm in section 3.1 to solve the relaxed LP. But there are three options of how to choose columns corresponding to the path variable x to be included in the initial relaxation at each node. We can take the following options:

(1) No path columns being included in the initial relaxation except the artificial variables, it is like the initial relaxation at root node;

(2) All path columns generated so far being included in the initial relaxation;

(3) Only a subset of all the path columns generated before being included in the initial relaxation. (If we take this option, we need to give a rule on selecting the path columns.)

In section 4, we give the computational results on some randomly generated instances.

4 Computational results

In table 3, 4 categories of problems are listed. For each problem category, 5 instances are randomly generated. In table 4, the initial sizes for these problems are listed.

problem category	which model	number of nodes	number of links	number of s-t pairs	number of instances
1m10	model 1	10	20	20	5
2m10	model 2	10	20	20	5
3m10	model 3	10	20	20	5
1m15	model 1	15	25	50	5
2m15	model 2	15	25	50	5
3m15	model 3	15	25	50	5

Table 3: Problem data

problem	link-based model			path-based model		
category	rows	columns	non-zeroes	rows	columns	non-zeroes
1m10	230	820	2460	50	40	80
2m10	240	840	2480	60	60	100
3m10	220	840	2440	40	60	60
1m15	790	2525	7575	90	75	125
2m15	800	2550	7600	100	100	150
3m15	775	2550	7550	75	100	100

Table 4: The initial sizes

Every instance is solved twice with MINTO (the Mixed INTeger Optimizer), once in link-based model, once in path-based model. For link-based model, we implement MINTO in default way, which is a kind of branch-and-cut algorithm, because some cuts may be generated by the default system functions including clique generation, implication generation, knapsack cover generation, flow cover generation. In the default way, MINTO will also do preprocessing and limited probing. For path-based model, the branch-and-price algorithm described in section 3 is applied with option (2) on selecting initial set of path columns at every node in the branch-tree. Since we want to do column generation to capitalize the special structure of the path-based model, we deactivate all the default system functions offered by MINTO, and program some application functions to implement column generation with MINTO.

In table 5, the average computational results over 5 instances for each problem category are displayed.

problem	link-based model			path-based model				
category	nodes	time t_l	cuts	nodes	time t_p	columns	t_l/t_p	gap
1m10	746	793.4	563	2315	221.2	2181	3.59	73.1%
2m10	2452	642.0	0	3445	634.8	3009	1.01	281.8%
3m10	1225	281.0	0	1614	101.6	980	2.77	183.4%
1m15	886	5650.6	1525	5402	1112.0	3925	5.08	62.1%
2m15	10684	20520.6	0	17905	12668.6	6688	1.62	115.2%
3m15	15121	25368.2	0	25688	8724.2	4942	2.91	93.0%

Table 5: Average computational results on 5 instances

nodes: average number of nodes evaluated in the branch-bound tree;
time : average CPU seconds (IBM RS/6000 model 550);
cuts: average number of cuts generated;
columns: average number of columns generated;
gap: average (IP-LP)/LP, where IP is the optimal solution, LP is the relaxed linear programming sulotion at the root node.

In table 5, we see that our branch-and-price algorithm on path-based models works better than the branch-and-cut algorithm on link-based model.

5 Algorithms on Model 2

In Model 1, if we change the flow variables f_{ij}^h to be binary, we have Model 2.

We will use the standard algorithm to solve the link-based model for Model 2, so here we will focus on how to solve the path-based model.

In Model (P_3) in Section 2, if we change path variables x_j^h to be binary, we have Model (P_4):

We will still use branch-and-price algorithm to solve (P_4) like we do on (P_3). But because we use column generation on path variables to solve the LP relaxation, we need to give

$$\min \sum_{e_m \in E} \sum_{r=1}^{T} k_m^r y_m^r$$

$$\begin{aligned}
\sum_{h \in H} \sum_{j=1}^{n_h} a_{mj}^h R_h x_j^h \quad &\le \quad \textstyle\sum_{r=1}^{T} c_r y_m^r \quad && \forall e_m \in E && (29)\\
\sum_{j=1}^{n_h} x_j^h \quad &= \quad 1 && \forall h \in H && (30)\\
y_m^r \ \ integer & && \forall e_m \in E, 1 \le r \le T && (31)\\
x_j^h \ \ binary & && \forall h \in H, 1 \le j \le n_h && (32)
\end{aligned}$$

Model P_4: Path-based Model with Binary x's

suitable branching strategy on x variables so that it works and the subproblem can be solved efficiently.

A typical branch rule on binary variables is that if x_j^h is fractional, two new branches are generated, one of these would force x_j^h to be 1, the other would force x_j^h to be 0. But with column generation, this typical branch rule does not work. In the branch with $x_j^h = 1$, there is no problem; But in the branch with $x_j^h = 0$, when we use column generation to solve the LP relaxation, that path variable x_j^h may be generated again, which causes problem. Therefore, we must give some branch rule on path variables such that if a path variable is set to 0 in some branch, it will be prevented to be generated again in that branch.

We give two branch rules which work, and will compare these two branch rules based on some computational results later.

Suppose we want to branch on path variable x_j^h, which represents a path from node s to node t, and x_j^h is composed of by links $e_1, e_2, ..., e_k$ in order from the origin to the destination. Instead of generating two branches, we generate k+1 branches.

Rule one:

(1) At the first branch, set $x_j^h = 1$;

(2) Force $x_j^h = 0$ by creating k new branches, one for each link e_i in the path such that in this branch no path from s to t is allowed to go through link e_i.

The subproblem under Rule one:

Suppose at current node in the branch tree, for demand $h \in H$, links $e_{h_1}, e_{h_2}, ..., e_{h_t}$ are not allowed, than the subproblem is:

For each demand $h \in H$, solve the shortest path problem in the network with links $e_{h_1}, e_{h_2}, ..., e_{h_t}$ deleted and with weight $w_m^* \ge 0$ on the rest links in the original network, where w_m^* is the dual variables as defined in Section 3.

Rule two:

(1) At the first branch, set $x_j^h = 1$;

(2) Force $x_j^h = 0$ by creating k new branches, one for each link e_i in the path such that in this branch any path from s to t must go through link e_j for j=1, 2, ..., i-1, but without going through e_i.

The subproblem under Rule two:

Suppose at current node in the branch tree, for demand $h \in H$, links $e_{h_1}, e_{h_2}, ..., e_{h_t}$ are fixed and $e_{h_{t+1}}$ is not allowed, and node u_h is adjacent to both node e_{h_t} and node $e_{h_{t+1}}$. Let s_h be the source node and t_h be the destination node for demand $h \in H$, then Rule two guarantee that $[e_{h_1}, e_{h_2}, ..., e_{h_t}]$ is in order a path from s_h to u_h, and the subproblem is:

For each demand $h \in H$, solve the from u_h to t_h shortest path problem in the network without nodes in the path $[e_{h_1}, e_{h_2}, ..., e_{h_t}]$ except node u_h and without link $e_{h_{t+1}}$, and with weight $w_m^* \geq 0$ on the rest links in the original network, where w_m^* is the dual variables as defined in Section 3. After the shortest path from u_h to t_h is found, the shortest path from s_h to t_h is found by add the length of the path $[e_{h_1}, e_{h_2}, ..., e_{h_t}]$ from s_h to u_h and the shortest path from u_h to t_h

Rule one and Rule two appear slightly different. But with Rule One, there are some overlaps between the k new branches created to force $x_j^h = 0$; With Rule Two, no overlaps happen.

It turns out that with the two branch rules on path variables, the subproblem for the column generation remains to be shortest path problem, which is a good feature for our algorithm.

6 Conclusions

From the computational results given in section 4, we see that the branch-and-price algorithm on the path-based model works better compared with the branch-and-cut algorithm on link-based model. But our initial inspiration on studying the path-based model was to study telecommunication network design problem with survivability constraints, in which some extra constraints, such as two-connected (for each s-t pair, two link disjoint paths exist) constraints, are added to the network design model. Because of the structure of the survivability constraints, it is suitable to set up a path-based model. One of our future research directions is to study how to incorporate the branch-and-price algorithm with extra survivability constraints in the model. The extra survivability constraints will make the subproblem for column generation more complicated. Another research direction is how to incorporate the branch-and-price algorithm with integral path variables. The requirement for these path variables to be integral is necessary for some practical problems, for example, if every commodity corresponding to a s-t pair can not be split. This will complicates the branching strategy and the subproblem for column generation.

BIBLIOGRAPHY

1. T. L. Magnanti and R. T. Wong (1984), "Network Design and Transportation planning: Models and Algorithms", Transportation Science 18, 1-55.

2. T. L. Magnanti, P. Mirchandani and R. Vachani (1991), "Modeling and Solving the Two Facility Capacitated Network Loading Problem", To appear in Opes. Res.

3. C. Barnhart, E. L. Johnson, G. L. Nemhauser, M.W.P. Savelsbergh and P. H. Vance, "Branch-and-Price: Column Generation for Solving Huge Integer Programs"

E-24-603
#4

Final Report on NSF Project on 'Design of Survivable Heterogeneous Data Networks'

NSF Grant NCR 92-06764

G. Anandalingam, University of Pennsylvania
Lloyd W. Clarke, Georgia Institute of Technology
January 1996

This is the final report on the above NSF grant. The work involves the development of formulations and novel solution techniques for designing heterogeneous data networks to survive under link or node failure.

We have graduated two Ph.D.s at the University of Pennsylvania, Keesung Nam who is now employed by A.T.& T. Bell Laboratories, and Frances Lee who has also received an offer from A.T.& T. Bell Labs. In March 1996, a third Ph.D., Panjing Gong, will be graduating from Georgia Institute of Technology.

We are completing a book on the design of telecommunications networks which will be published by Kluwer. Three papers from our research will shortly appear in technical journals, and three other papers are currently under review. We have also made a number of presentations of our results in international conferences and academic and professional organizations.

Based on our international reputation in telecommunications network design, Anandalingam has been invited to be a Visiting Lecturer in the Systems Analysis Laboratory of the Helsinki University of Technology (Finland), and to be the lead person in creating a Center for Information Technology and Telecommunications at the Indian Institute of Management in Bangalore, a city which is now being called the "Silicon Valley" of India.

We will briefly describe our research on the NSF grant below:

1 Survivable Network Algorithms

We have derived methodologies for designing survivable networks. The work involves formulating the survivable network design problem, obtaining methodologies for solving problems with numerous nodes, integrating these heuristics into a Design System, and obtaining numerical results that show the computational performance of the Design System.

Our research in the area of survivable networks has included both heuristics and mathematical programming approaches. We have developed a software package that designs both capacitated and uncapacitated networks. The heuristic portion of the system works in two steps. The first step is the design of an initial feasible survivable network. The second step is the improvement of this solution. We have shown that this system generates very good solutions. The quality of the heuristic solutions can be assessed by the math programming portion of the design system. The math programming portion can determine lower bounds for both capacitated and uncapacitated networks. It can also find optimal solutions for uncapacitated survivable networks.

Some of this work has been written for publication. The paper "A Bootstrap Heuristic for Designing Minimum Cost Survivable Networks" (L. Clarke and G. Anandalingam) has been accepted for publication in the journal *Computers and Operations Research.* This paper deals with uncapacitated survivable networks. It covers mathematical formulations for link and node survivability. It also uses these formulations to determine lower bounds and the optimal solution. We present a heuristic based on the lower bound that generates an initial feasible solution. We use several different improvement heuristics to generate a final solution.

Another paper entitled "An Integrated System for Designing Minimum Cost Survivable Telecommunications Networks" (L. Clarke and G. Anandalingam) is scheduled to appear in *IEEE Journal on Systems, Man & Cybernetics.* This paper concentrates on capacitated survivable networks. A Lagrangian relaxation technique is used for finding lower bounds. Several heuristics are presented for finding solutions.

We also have a paper "The Design of Minimum Cost Survivable Telecommunication Networks: A Synthesis." (L. Clarke and G. Anandalingam). This is a survey paper that covers the difference between network reliability and network survivability. Fundamental concepts such as disjoint paths and biconnected components are covered. We present formulations and sample research using link-based designs and path-based designs. The paper has been circulated for comments and will be submitted to a journal shortly.

General Network Algorithms: In the search for more efficient methods of solving network design problems, we looked at new formulations for the problem of designing survivable telecommunications networks. We make the distinction between network design based on links where one has to decide what *links* to place in the network to ensure survivability, and design based on *paths* where one has to choose among a set of paths for each origin-destination pair to ensure survivability. The path based approach has the advantage of being able to deal better with capacity constraints, origin destination flow requirements, and delay constraints. The examination of general algorithms for survivable network design is the area of dissertation research for Panjing Gong of Georgia Institute of Technology.

The path based models require a completely different mathematical approach. To better develop this approach, we return to the basic network design problem without survivability constraints. In these models, the number of variables is a factorial function. A 10 node network can have over 2 million variables. This is handled by using column generation techniques, where the initial problem will only have a small subset of the variables and additional variables are added as needed as the algorithm progresses. Determining how and when to add variables to the model requires a reasonable amount of optimization theory. This theory isn't easily applied to survivable networks. The emphasis is general networks with capacity constraints.

We worked with three different models. The first model is a simple capacitated model with choice of link locations. The second model has multiple size links available. Each location can have only one of the several different size links. The final model has multiple link sizes and can install any number of links in any location. Each of these models have a split flow and a unified flow version. To work with the requirement that origin-destination flow must remain on one path. required some innovative mathematical thought that relates to the column generation in branch and bound. We have found that the path formulation method is a more efficient and very promising approach to solving network design problems. Optimal solutions can be found for problems in size of up to 20 nodes. Comparing our path based formulation to a standard link based formulation we can find optimal solutions up to 5 times faster.

We have presented this research in its different stages at different technical conferences (See Appendix 2).

2 Design of Survivable International Networks

This research is the basis of the dissertation of Keesung Nam, a Ph.D. student supervised by G. Anandalingam at the University of Pennsylvania entitled *The Design of Survivable International Telecommunications Networks*. The main differences between designing survivable data networks for corporations located in one country, and distributed across the globe are: (a) the peak communications traffic between pairs of countries occurs at different times, and (b) different countries have different objectives, although they may use the same network. Thus, telecommunications networks can be used very efficiently by implementing dynamic routing that sends traffic via transit countries.

In Nam's dissertation, we have formulated two kinds of problems. First, we have examined a survivable network that uses pre-planned time varying dynamic routing based on the daily traffic profile among countries (Chapter 3). Each country-to-country pair can choose the degree of survivability by giving the network designer a parameter that fixes the amount of traffic that should be re-routed under node or link failure. We have also derived heuristics to solve large size problems and have tested the heuristics against lower bounds. The heuristic is very simple: It involves decomposing the problem, using the Design System of Clarke and Anandalingam to solve part of the new structure, and in the other part, using linear programming to solve integer programming problems and rounding the solutions in a pre-arranged order. We have tested the problem on both an Atlantic and a Pacific Basin network. The Pacific Basin network is large (containing 15 countries) with more than 40,000 variables, many of them integer, and around 60,000 constraints. The problem was solved in 16-33 minutes on a SUN SparcStation, and was within 1 per cent of the Lower Bound. These are excellent results. A paper entitled 'Heuristics for Designing International Telecommunications Networks', co-authored by Nam, Anandalingam and Clarke has been submitted for publication.

The second problem we examined involved the advantages and disadvantages for the countries to form a coalition in order to design a truly integrated network. The issues were the increase in benefits and decrease in costs to the entire coalition of designing the network cooperatively, and how these benefits should be distributed to the players equitably. (Chapter 5 in Nam's dissertation). For this we used developments in Co-operative Game Theory. We also wanted to examine whether the U. S. should take the lead in developing the international network. The issue here is whether there are economic gains to be made by being the leader in a non-cooperative Stackelberg game, or whether one should only be part of a coalition. To analyze this case, we had to develop some methodology in bi-level programming (Chapter 4). All of the numerical analysis in these cases were based on real world data that was distilled very carefully from published sources. A paper entitled "Conflict and Cooperation in Designing International Telecommunications Networks" co-authored by Anandalingam and Nam has been submitted for publication.

3 The LAN-LAN Internetworking Problem

This area of research is the basis of the Ph.D. dissertation of Frances Lee who is supervised by G. Anandalingam at the University of Pennsylvania. This work will continue on to next year, and we expect that the dissertation will be completed by May 1995.

We have examined a number of bridges and routers in the commercial market, and have extracted out all the important features. The issue is which device one would use in interconnecting different LANs. This examination has resulted in a draft of a paper entitled "On the Choice of Bridges and Routers as Internetworking Devices" that has been submitted for publication.

We have also modeled the survivable LAN-LAN internetworking problem using a hierarchical architecture. In this case, we use either a combination of remote routers and backbone routers or bridges and backbone routers. The remote network is designed using both a mathematical programming approach and a statistical clustering approach. We have examined the computational performance of each. The backbone network is designed using the Design System that was produced by Clarke and Anandalingam, described earlier. We also have a heuristic for connecting the remote and the backbone networks. The methodology has been applied to the real world case of designing a LAN-LAN network for the Penn campus. Using our methodology, we can design survivable LAN-LAN internetworks with 500 users in less than 3-5 minutes which is extremely fast. We have submitted a paper entitled "Design of Survivable LAN-LAN Internetworks with a Hierarchical Topology", co-authored by Lee and Anandalingam that will appear in *IEEE Transactions on Networking.*

We have examined the performance of the internetworks that we have designed using the heuristics described earlier relative to networks with different topologies and devices. In particular, we have examined networks with routers instead of bridges, and a combination of bridges in remote sites and routers on the backbone and situations where the backbone is a ring versus where the access network is a ring. In addition, we have also examined different dynamic routing mechanisms in the network. All of these analyses involved using discrete event simulation. We are in the process of finalizing a paper (which will be the final chapter in Frances Lee's Ph.D. dissertation) which will also be submitted for publication.

4 Availability of Self-Healing Ring Networks

Our collaboration with Mark Wilson of AT&T Bell Laboratories on estimating the availability of self healing ring networks was quite fruitful. He completed his dissertation in May 1995. Although Mark Wilson is *not* being funded by the NSF grant, his dissertation is intimately connected to the subject of survivable heterogeneous networks.

Mark Wilson examined models for evaluating the availability of networks that are supposed to be survivable (or self healing). While much of our previous work has been on providing a survivable topology, Wilson's dissertation was on evaluating the probabilities of such networks actually being survivable at points in time and after some failure event (i.e. the 'availability' of such networks). The focus of this work is on SONET (Synchronous Optical Network) which is a survivable network because of its ring topology, and 'availability' will be one of components that customers will use (along with cost, functionality etc) in order to assess their networks.

The collaboration involved the development of heuristics to estimate availability rapidly. In addition, these availability models are being calibrated on SONET, and should be used in real world applications. G. Anandalingam was a member of Mark Wilson's dissertation committee. We expect to have 1-2 scholarly papers from the dissertation.

Appendix 1

Research Output

A. Ph.D. Dissertations

- Keesung Nam, *The Design of International Telecommunications Networks*, Ph.D., University of Pennsylvania, December 1994.
- Frances Lee, *Design of Survivable Heterogeneous Telecommunications Networks*, Ph. D., University of Pennsylvania, December 1995.
- Panjing Gong, *Capacitated Network Design with Column Generation*, Ph.D., Georgia Institute of Technology, Expected March 1996.

B. Publications

- Lloyd W. Clarke and G. Anandalingam. "A bootstrap heuristic for designing minimum cost survivable networks", *Computers and Operations Research*. vol. 22. no. 9. pp 921-934. 1995
- Lloyd W. Clarke and G. Anandalingam. "An integrated system for designing minimum cost survivable telecommunications networks", *IEEE Transactions on Systems. Man. and Cybernetics*, vol. 26, no. 9. 1996 (in press).
- Frances Lee, G. Anandalingam, and Lloyd W. Clarke, "Design of LAN-LAN internetworks with a hierarchical topology", *IEEE/ACM Transactions on Networking*. 1996 (forthcoming).
- G. Anandalingam, *Design of Large Scale Telecommunications Networks*. Kluwer Scientific Publishers, Norwood, MA, 1996 (forthcoming).

C. Currently Under Review

- G. Anandalingam and Keesung Nam, "Conflict and cooperation in designing international telecommunications networks".
- Keesung Nam, G. Anandalingam, and Lloyd W. Clarke, "Heuristics for designing survivable international telecommunications networks".

- Frances Lee and G. Anandalingam, "On the choice of bridges and routers as internetworking devices".

- G. Anandalingam and Frances Lee, "A Statistical Heuristic for Access Telecommunications Network Design"

D. Working Papers

In addition to those listed under B & C above.

- Lloyd W. Clarke and G. Anandalingam, "The Design of Minimum Cost Survivable Telecommunication Networks: A Synthesis", Working Paper LEC94-12, School of Industrial and Systems Engineering, Georgia Institute of Technology, Atlanta, GA 30332.

- P. Gong, Lloyd W. Clarke, and G. Anandalingam, "Capacitated Network Design with Column Generation", Working Paper, LEC95-08, School of Industrial and Systems Engineering, Georgia Institute of Technology, Atlanta, GA 30332.

Appendix 2

Presentations

- G. Anandalingam, Lloyd W. Clarke and Frances Lee, 'Heuristics for Designing Survivable Heterogeneous Telecommunications Networks', INFORMS Conference on the Interface of Computers and Operations Research, Dallas, January 5, 1996.
- G. Anandalingam, 'On the Design of Telecommunications Networks: Do Sophisticated Models Buy One Anything?", Dept of Systems Engineering, University of Virginia, October 13, 1995.
- G. Anandalingam, 'New Approaches for Designing Large Scale Heterogeneous Telecommunications Networks', A T & T Bell Laboratories, September 19, 1995.
- G. Anandalingam and Keesung Nam, 'Design of Survivable International Telecommunications Networks', TIMS 33rd International Conference, Singapore, June 27, 1995.
- G. Anandalingam, 'The Design of Heterogeneous Telecommunications Networks: To Anneal or Not to Anneal', Dept of Civil Engineering and Operations Research, Princeton University, March 9, 1995.
- Lloyd W. Clarke and P. Gong, "An Algorithm for Designing Networks with Column Generation and Branch and Bound," presented at ORSA/TIMS Joint National Meeting Pheonix, November 1993.
- Lloyd W. Clarke and P. Gong, "Network Design with Column Generation" presented at Mathematical Programming Symposium, Ann Arbor, August, 1994.
- Lloyd W. Clarke and P. Gong, "A Path-Based Formulation for Designing Telecommunication Networks," presented at ORSA/TIMS Joint National Meeting, Detroit, October 1994.
- G. Anandalingam, 'Global Communications: Present Status and Future Trends', Computer Society of India, Bangalore, India, August 3, 1993.